

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 16:09:27 ; Search time 299.73 Seconds
(without alignments)
18.379 Million cell updates/sec

Title: US-09-331-631a-7_COPY_34_80
Perfect score: 258
Sequence: 1 YERDPQQYEQCQRCSEAA.....QCEQFCERYKEQQQQEE 47

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

listing first 45 summaries

Database : SPTRMBL_15,*

- 1: sp_archea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rabbit:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match length	DB ID	Description
1	258	100.0	525	10	Q43358 theobroma cacao
2	157	60.9	625	10	Q9SP13 macadamia integrifolia
3	157	60.9	666	10	Q9SP14 macadamia integrifolia
4	154	59.7	666	10	Q9SP15 macadamia integrifolia
5	146	56.6	593	10	Q9SEW4 juglans regia
6	95	36.8	1038	5	Q60983 dictyostelium discoideum
7	90.5	35.1	1089	12	Q40947 kaposi's sarcoma-associated herpesvirus
8	90	34.9	810	10	Q9ZW13 cucurbita maxima
9	89.5	34.7	1129	12	Q9QR71 kaposi's sarcoma-associated herpesvirus
10	88.5	34.3	393	10	Q9ZP00 oryza sativa
11	85	32.9	551	10	Q43607 prunus dulcis
12	85	32.9	1162	12	Q98148 kaposi's sarcoma-associated herpesvirus
13	84	32.6	1108	5	Q9ND10 babesia microti
14	84	32.6	1339	11	Q35788 rattus norvegicus
15	84	32.6	2123	5	Q9U957 dictyostelium discoideum
16	83.5	32.4	251	4	Q9NW29 homo sapiens
17	83	32.2	919	4	Q9ngs7 homo sapiens
18	82.5	32.0	554	5	Q9VPS3 drosophila melanogaster
19	82	31.8	314	5	Q9V26 drosophila melanogaster

ALIGNMENTS

RESULT	1	PRELIMINARY;	PRN;	525 AA.
ID	Q43358			
AC	Q43358;			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)			
DE	VICILIN PRECURSOR.			
GN	Theobroma cacao (cacao).			
OS	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Magnoliophyta; eudicots; core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Theobroma.			
OC	OC			
OX	NCBI_TaxID=3641;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=LEAVES;			
RX	MEDLINE=92288309; PubMed=1600151;			
RA	McHenry L., Fitz P.J.;			
RT	Comparison of the structure and nucleotide sequences of vicilin genes of cocoa and cotton raise questions about vicilin evolution.;			
RL	Plant Mol. Biol. 18:1173-1176(1992).			
DR	EMBL; X62625; CAA44493; 1; -.			
DR	HSSP; P02853; 2PHI.			
DR	MENDEL; 30919; TMCC; 1188; 30919.			
DR	INTERPRO; IPR001113; -.			
DR	PFAM; PF00546; Seedstore_7s; 1.			
DR	PRODOM; PD001059; -; 1.			
DR	KW SIGNAL.			
FT	Q9ZW13 cucurbita maxima			
FT	Q9QR71 kaposi's sarcoma-associated herpesvirus			
FT	Q9ZP00 oryza sativa			
FT	Q43607 prunus dulcis			
FT	Q98148 kaposi's sarcoma-associated herpesvirus			
FT	Q9ND10 babesia microti			
FT	Q35788 rattus norvegicus			
FT	Q9U957 dictyostelium discoideum			
SEQUENCE	525 AA:	60798 MW:	19114CD5C248905D CRC64:	
QY	1 YERDPQQYEQCQRCSEAA.....QCEQFCERYKEQQQQEE 47			
DB	34 YERDPQQYEQCQRCSEAA.....QCEQFCERYKEQQQQEE 80			

Query Match 100.0%; Score 258; DB 10; Length 525;
Best local Similarity 100.0%; Pred. No. 5.3e-20;
Matches 47; Conservative 0; Mismatches 0;
Indels 0; Gaps 0;

Query Match 100.0%; Score 258; DB 10; Length 525;
Best local Similarity 100.0%; Pred. No. 5.3e-20;
Matches 47; Conservative 0; Mismatches 0;
Indels 0; Gaps 0;

RESULT	2	RESULT	4
ID	O9SP3	ID	O9SP5
AC	O9SP3;	AC	O9SP5;
DT	01-MAY-2000 (TREMBLrel. 13, Created)	DT	01-MAY-2000 (TREMBLrel. 13, Created)
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)	DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)	DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE	VICILIN PRECURSOR (FRAGMENT).	DE	VICILIN PRECURSOR.
GN	AMP2.	GN	AMP2.
OS	Macadamia integrifolia (Macadamia nut).	OS	Macadamia integrifolia (Macadamia nut).
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;	OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC	Magnoliophyta; eudicots; Proteaceae; Macadamia.	OC	Magnoliophyta; eudicots; Proteaceae; Macadamia.
NCBI_TaxID=60698;		NCBI_TaxID=60698;	
[1]		[1]	
RP	SEQUENCE FROM N_A.	RP	SEQUENCE FROM N_A.
RC	TISSUE=NUT KERNEL;	RC	TISSUE=NUT KERNEL;
RA	Marcus J.P., Goultier K.C., Green J.L., Manners J.M.;	RA	Marcus J.P., Goultier K.C., Green J.L., Manners J.M.;
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia.";	RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia kernels.";
RL	Plant J. 0:0-0 (1999).	RL	Plant J. 0:0-0 (1999).
EMBL	AF161885; AAD54246.1; -.	EMBL	AF161885; AAD54244.1; -.
HSSP	P02833; 2PHL.	HSSP	P02833; 2PHL.
DR	INTERPRO; IPR001113; -.	DR	INTERPRO; IPR001113; -.
DR	PFAM; PF00546; Seedstore_7s; 1.	DR	PFAM; PF00546; Seedstore_7s; 1.
FT	NON_TER 1	FT	NON_TER 1
SQ	SEQUENCE 625 AA; 73586 MW; 415808A89D370296 CRC64;	SQ	SEQUENCE 666 AA; 78217 MW; C752B884B2DF0224 CRC64;
RESULT	3	Query Match	59.7%
O9SP4		Best Local Similarity	50.9%
ID	O9SP4	Matches	53.3%
AC	O9SP4;	Conservative	53.3%
DT	01-MAY-2000 (TREMBLrel. 13, Created)	Mismatches	8;
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)	Indels	0;
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)	Gaps	0;
DE	VICILIN PRECURSOR.		
GN	AMP2.		
OS	Macadamia integrifolia (Macadamia nut).		
OC	Eukaryota; Viridiplantae; Embryophyta; Spermatophyta;		
OC	Magnoliophyta; eudicots; Proteaceae; Macadamia.		
NCBI_TaxID=60698;			
[1]			
RP	SEQUENCE FROM N_A.		
RC	TISSUE=NUT KERNEL;		
RA	Marcus J.P., Goultier K.C., Green J.L., Manners J.M.;		
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia kernels.";		
RL	Plant J. 0:0-0 (1999).		
EMBL	AF161885; AAD54245.1; -.		
HSSP	P02833; 2PHL.		
DR	INTERPRO; IPR001113; -.		
DR	PFAM; PF00546; Seedstore_7s; 1.		
FT	NON_TER 1		
SQ	SEQUENCE 666 AA; 78243 MW; 0ECA22F8710F8A7B CRC64;		
Query Match	60.9%	Score	157;
Best Local Similarity	53.3%	DB	10;
Matches	53.3%	Length	666;
24;		Pred.	No. 2.8e-09;
Conservative	12;	Mismatches	9;
Indels	0;	Indels	0;
Gaps	0;	Gaps	0;
RESULT	3	Query Match	59.7%
O9SEW4		Best Local Similarity	51.1%
ID	O9SEW4	Matches	23;
AC	O9SEW4;	Conservative	13;
DT	01-MAY-2000 (TREMBLrel. 13, Created)	Mismatches	9;
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)	Indels	0;
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)	Gaps	0;
DE	VICILIN-LIKE PROTEIN PRECURSOR (FRAGMENT).		
GN	Eukaryota; Virgiliplanta; Embryophyta; Spermatophyta;		
OS	Juglans regia (English walnut).		
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;		
OC	Magnoliophyta; eudicots; core eudicots; Rosidae; eurosids 1;		
OC	Fagales; Juglandaceae; Juglans.		
NCBI_TaxID=51240;			
[1]			
RP	SEQUENCE FROM N_A.		
RC	STRAIN=CV; SUNLAND; TISSUE=SOMATIC EMBRYO LINE;		
RA	Teuber S.S., Jarvis K.C., Peterson W.R., Danodekar A.M., Anvari A.A.;		
RT	"Identification and cloning of a cDNA encoding a vicilin-like protein, Jug r 2, from English walnut kernel (Juglans regia): a major food allergen.";		
RT	Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.		
RL	EMBL; AF06055; AAF18269.1; -.		
DR	HSSP; P02833; 2PHL.		
DR	INTERPRO; IPR001113; -.		
DR	PFAM; PF00546; Seedstore_7s; 1.		
FT	NON_TER 1		
SQ	SEQUENCE 593 AA; 6990 MW; 9BA127E19B18C0D8 CRC64;		
Query Match	56.6%	Score	146;
Best Local Similarity	45.8%	DB	10;
Matches	27;	Length	593;
Conservative	8;	Pred.	No. 3.7e-08;
Mismatches	10;	Indels	14;
Indels	14;	Gaps	1;
Gaps	1;		
RESULT	4	Query Match	56.6%
O9SP5		Best Local Similarity	45.8%
ID	O9SP5	Matches	27;
AC	O9SP5;	Conservative	8;
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)	Mismatches	10;
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)	Indels	14;
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)	Gaps	1;
DE	VICILIN PRECURSOR.		
GN	AMP2.		
OS	Macadamia integrifolia (Macadamia nut).		
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;		
OC	Magnoliophyta; eudicots; Proteaceae; Macadamia.		
NCBI_TaxID=60698;			
[1]			
RP	SEQUENCE FROM N_A.		
RC	TISSUE=NUT KERNEL;		
RA	Marcus J.P., Goultier K.C., Green J.L., Manners J.M.;		
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia.";		
RL	Plant J. 0:0-0 (1999).		
EMBL	AF161884; AAD54245.1; -.		
HSSP	P02833; 2PHL.		
DR	INTERPRO; IPR001113; -.		
DR	PFAM; PF00546; Seedstore_7s; 1.		
FT	NON_TER 1		
SQ	SEQUENCE 593 AA; 6990 MW; 9BA127E19B18C0D8 CRC64;		
Query Match	56.6%	Score	146;
Best Local Similarity	45.8%	DB	10;
Matches	27;	Length	593;
Conservative	8;	Pred.	No. 3.7e-08;
Mismatches	10;	Indels	14;
Indels	14;	Gaps	1;
Gaps	1;		
RESULT	4	Query Match	56.6%
O9SP5		Best Local Similarity	45.8%
ID	O9SP5	Matches	27;
AC	O9SP5;	Conservative	8;
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)	Mismatches	10;
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)	Indels	14;
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)	Gaps	1;
DE	VICILIN PRECURSOR.		
GN	AMP2.		
OS	Macadamia integrifolia (Macadamia nut).		
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;		
OC	Magnoliophyta; eudicots; Proteaceae; Macadamia.		
NCBI_TaxID=60698;			
[1]			
RP	SEQUENCE FROM N_A.		
RC	TISSUE=NUT KERNEL;		
RA	Marcus J.P., Goultier K.C., Green J.L., Manners J.M.;		
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia kernels.";		
RL	Plant J. 0:0-0 (1999).		
EMBL	AF161885; AAD54244.1; -.		
HSSP	P02833; 2PHL.		
DR	INTERPRO; IPR001113; -.		
DR	PFAM; PF00546; Seedstore_7s; 1.		
FT	NON_TER 1		
SQ	SEQUENCE 593 AA; 6990 MW; 9BA127E19B18C0D8 CRC64;		
Query Match	56.6%	Score	146;
Best Local Similarity	45.8%	DB	10;
Matches	27;	Length	593;
Conservative	8;	Pred.	No. 3.7e-08;
Mismatches	10;	Indels	14;
Indels	14;	Gaps	1;
Gaps	1;		

Db	117 QDPQQQYHRCCRCQ1QFOSPERORQCRCRQKEQQGRGPEASPRRRESRGREEE 175	QY	2 ERDPROQECCRCEATTEPQEQCEREKQQQQEE 47
RESULT	6	PRELIMINARY;	PRT; 1038 AA.
ID	060983		
AC	060983;		
DT	01-AUG-1998 (TREMBREL 07, Created)	RESULT	8
DT	01-AUG-1998 (TREMBREL 07, Last sequence update)	ID	092W13
DT	01-MAY-2000 (TREMBREL 13, Last annotation update)	ID	092W13
DE	REP-LIKE.	AC	092W13;
GN	REP.	DT	01-MAY-1999 (TREMBREL 10, Last sequence update)
OS	Dictyostelium discoideum (Slime mold).	DT	01-OCT-2000 (TREMBREL 15, Last annotation update)
OG	Plasmid Ddp5.	DE	PV100.
OC	Eukaryota; Dictyostellida; Dictyostelium.	OS	Cucurbita maxima (Pumpkin) (Winter squash).
OX	NCBI_TaxID=44689;	OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Magnoliophyta; eudicots; Rosidae; eurosids I;
RN	[1]	OC	Cucurbitales; Cucurbitaceae; Cucurbita.
RP	SEQUENCE FROM N.A.	OX	NCBI_TaxID=3661;
RC	STRAIN=N52162;	RN	[1]
RX	MEDLINE=98198836; PubMed=9539429;	RP	SEQUENCE FROM N.A.
RA	Rieben W.K. Jr., Gonzales C.M., Gonzales S.T., Pilkington K.J.,	RC	STRAIN=NKUROKAWA_AMAKURI_NANKIN; TISSUE=COTYLEBON;
RA	Kiyosawa H., Hughes J.E., Welker D.L.,	RX	MEDLINE=99107919; PubMed=891029;
RT	"Dictyostelium discoideum nuclear plasmid Ddp5 is a chimera related to the Ddp1 and Ddp2 plasmid families.";	RA	Yamada K., Shimada T., Kondo M., Nishimura M., Hara-Nishimura I.;
RT	Genetics 148:1117-1125(1998). [2]	RT	Multiple functional proteins are produced by cleaving Asn-Gln bonds of a single precursor by vacuolar processing enzyme.;"
RL	RN	RL	J. Biol. Chem. 274:2563-2570(1999).
RP	SEQUENCE FROM N.A.	DR	EMBL; AB019195; BA3A4056.1; -.
RC	STRAIN=WS2162;	DR	HSSP; P02853; 2PIPL.
RA	Rieben W.K., Gonzales C., Gonzales S.T., Pilkington K., Kiyosawa H.,	DR	INTERPRO; IPR001113; -.
RA	Hughes J.E., Welker D.L.;	DR	PFAM; PF00546; Seedstore_7s; 1.
DR	Submitted (ABR-1997) to the EMBL/GenBank/DDBJ databases.	SQ	PRODOM; PD081059; -; 1.
KW	Plasmid.	SEQUENCE	810 AA; 97314 MW; A829A3F7542266AB CRC64;
SQ			
Query Match	36.8%; Score 95; DB 5; Length 1038;	Query Match	34.9%; Score 90; DB 10; Length 810;
Best Local Similarity	37.0%; Pred. No. 0.014; Mismatches 17; Indels 18; Gaps 0;	Best Local Similarity	34.9%; Pred. No. 0.038; Mismatches 17; Indels 15; Gaps 0;
Matches		Matches	
QY	2 ERDPROQECCRCEATTEPQEQCEREKQQQQEE 47	QY	5 PROYEQCOCRCESEATEERDQEQQCERCEKQQQQEE 47
Db	966 QQEQQQEQQQEQQQEQQQEQQQEQQQEQQQEQ 1011	Db	74 PRAEYEVCRKQVAERGVEQKRCQVCERLRRREQGED 116
RESULT	7	RESULT	9
ID	090947	Q9QR71	PRELIMINARY;
ID	090947	ID	PRT; 1089 AA.
AC	090947;	AC	
DT	01-JAN-1998 (TREMBREL 05, Created)	DT	01-MAY-2000 (TREMBREL 13, Last sequence update)
DT	01-JAN-1998 (TREMBREL 05, Last sequence update)	DT	01-MAY-2000 (TREMBREL 13, Last sequence update)
DT	01-JUN-2000 (TREMBREL 14, Last annotation update)	DE	LATENT NUCLEAR ANTIGEN.
DE	ORF 73.	GN	ORF 73.
OS	Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).	OS	Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;	OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC	Gammaherpesvirinae; Rhadinovirus.	OC	Gammaherpesvirinae; Rhadinovirus.
OX	NCBI_TaxID=37296;	OX	NCBI_TaxID=37295;
RN	[1]	RN	[1]
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
RX	MEDLINE=97296220; PubMed=9151804;	RC	STRAIN=GK18;
RA	Nelipel F., Albrecht J.C., Fleckenstein B.,	RA	Glenn M., Rainbow L., Aurade F., Davison A., Schulz T.F.;
RT	Cell-homologous genes in the Kaposi's sarcoma-associated rhadinovirus human herpesvirus 8: determinants of its pathogenicity?";	RT	"Identification of a Spliced Gene from Kaposi's Sarcoma-Associated
RT	J. Virol. 71:4187-4192(1997).	RT	Herpesvirus Encoding a Protein with Similarities to Latent Membrane Proteins 1 and 2A of Epstein-Barr Virus.";
RL	EMBL; 093872; AAB66501.1; -.	RL	J. Virol. 73:6953-6963(1999).
DR	INTERPRO; IPR002017; -.	DR	EMBL; AF148805; AAD66501.1; -.
DR	INTERPRO; IPR002017; -.	DR	INTERPRO; IPR00533; -.
DR	INTERPRO; IPR002965; -.	DR	INTERPRO; IPR002965;
DR	PRINTS; PRO0194; TROPOMYOSIN.	DR	PRINTS; PRO1217; PRICHEXBNS.
DR	SEQUENCE	SEQUENCE	1129 AA; 131346 MW; 8F63855B45F79109 CRC64;
Query Match	35.1%; Score 90.5; DB 12; Length 1089;	Best Local Similarity	41.3%; Pred. No. 0.043; Mismatches 15; Indels 1; Gaps 1;
Matches	19;	Matches	

